Old Document Restoration using Super Resolution GAN and Semantic Image Inpainting

Yong Jun Kim
Jeju National University
102, Jejudaehak-ro, Jeju-si,
Jeju-do, Republic of Korea
+82-64-754-3657
yjkim0123@gmail.com

DebapriyaHazra
Jeju National University
102, Jejudaehak-ro, Jeju-si,
Jeju-do, Republic of Korea
+82-64-754-3657
debapriyah@gmail.com

Yungcheol Byun
Jeju National University
102, Jejudaehak-ro, Jeju-si,
Jeju-do, Republic of Korea
+82-64-754-3657
ycb@jejunu.ac.kr

Khi-Jung Ahn
Jeju National University
102, Jejudaehak-ro, Jeju-si,
Jeju-do, Republic of Korea
+82-64-754-3652
(Corresponding Author)
kjahn@jejunu.ac.kr

ABSTRACT
Restoration of damaged images is a fundamental problem that has been attempted before the advent of digital image processing technology. In this paper, one of the deep neural network technologies (GAN), we propose an image restoration network using Generative Adversarial Network. The proposed system is the image generation network, the generation result plateIt consists of a star network. Old documents not only contain information, but also we can learn about historical people's thought and consciousness from the past. Old document restoration is referred to as the restoration of documents that are usually made of parchment which are damaged either naturally or artificially. Missing regions in old documents are filled based on the current visual data which is a hard task in image inpainting. In this paper, we present Super Resolution Generative Adversarial Network (SRGAN) and semantic image inpainting for restoring the old documents so that they can be reused.

CCS Concepts
• Computing methodologies→Image segmentation

Keywords
Old document restoration; SRGAN; Semantic Image Inpainting

1. INTRODUCTION
The image is damaged due to various causes. In the case of general photographs, the photo paper is damaged during the deterioration and storage of photo paper, but in the case of digital images, it is mainly caused by noise due to shooting conditions. In the restoration of such a damaged image, a narrow area problem can be corrected in the form of interpolation by using close information such as super-resolution or noise reduction method. However, in the case of large-scale damage in a large area, it is very difficult to solve this problem because it is necessary to reason about the lost area. With the recent introduction of deep neural networks (DNNs), remarkable developments are being made in various fields such as image recognition and picture quality improvement. Recently, a study related to encoding and decoding (autoencoder) in reconstructing lost images has been introduced. Later, this encoding-decoding scheme was applied to deep neural networks, and recently, a study was proposed to reconstruct an image by considering the global/local context of the image through a hostile generating neural network. The image generated by the hostile generating neural network is similarly generated according to the characteristics of the image to be generated, so it is generated in a structurally natural form.[1]

Restoring antique pictures and old documents is an emerging field in today's world. To retrieve essential knowledge from the past is a crucial and necessary task to obtain, extract and generate information for present or future. Machine Learning is a field of Artificial Intelligence which is used enormously in developing algorithms that can be used to generate new images and restore damaged digitized documents. Historical documents are often used or referred for inventing new ideas or proposing new methodologies in every field. Therefore, there is a need to rectify and find some automatic approach to refurbish the documents.

Ian Goodfellow in 2014 developed a Generative Adversarial Network (GAN) as shown in Fig. 1, which consists of two networks known as the generator and the discriminator which competes against each other in a zero-sum game to produce the new and best results based on the training set. Super Resolution GANs are used to generate high resolution images. Image inpainting is a process that helps in reconstructing any image whose quality has deteriorated and data has been lost. Inpainting based on the available data helps in generating the best version of the image. In our paper, we have used both super resolution (SR) GAN and image inpainting to reconstruct old damaged documents. We have trained SRGAN with 1000 images and then implied semantic image inpainting on the output images generated from SRGAN to rectify and restore the images.

WAIE 2019, November 25–27, 2019, Singapore, Singapore

ACM ISBN 978-1-4503-7623-5/19/11…$15.00
DOI: https://doi.org/10.1145/3397453.3397459

There are a variety of GAN's available that does different task. The four most commonly used generative adversarial networks are mentioned in Table 1 and Table 2 along with their loss function [2]. The distance between the distributions of the data generated by GAN and the distribution of the real data is termed as loss functions [3].

Table 1.Discriminator loss for GANs [1]

Type	Discriminator Loss
Generative Adversarial Network	$L_D^{GAN} = -E_{x\sim Pdata}[\log(D(x))] - E_{z\sim Pz}[\log(1 - D(G(z)))]$
Conditional Generative Adversarial Network	$L_D^{CGAN} = -E_{x\sim Pdata}[\log(D(x,c))] - E_{z\sim Pz}[\log(1 - D(G(z),c))]$
Info Generative Adversarial Network	$L_D^{InfoGAN} = L_D^{GAN} - \lambda L_1(c,c')$
Auxiliary Classifier Generative Adversarial Network	$L_D^{ACGAN} = L_D^{GAN} + E_{x\sim Pdata}[P(class = c\mid x)] + E_{z\sim Pz}[P(class = c\mid G(z))]$

Table 2. Generator loss for GANs [1]

Type	Generator Loss
Generative Adversarial Network	$L_G^{GAN} = -E_{x\sim Pz}[\log(D(G(z)))]$
Conditional Generative Adversarial Network	$L_G^{CGAN} = -E_{x\sim Pz}[\log(D(G(z),c))]$
Info Generative Adversarial Network	$L_G^{InfoGAN} = L_G^{GAN} - \lambda L_1(c,c')$
Auxiliary Classifier Generative Adversarial Network	$L_G^{ACGAN} = L_G^{GAN} + E_{z\sim Pz}[P(class = c\mid G(z))]$

Through competition of neural network models, they learn and produce results. The two models are called 'generators' and 'discriminators', with the opposite purpose. The constructor learns the actual data and generates false data based on it. The goal is to generate false data that is close to reality. The discriminator learns to determine whether the data produced by the constructor is real or false. The goal is not to play around with false data in the constructor. Ian Goodfellow likened the constructor to a counterfeit bill and the discriminator to the police. The constructor learns the data that did not deceive the discriminator, and the discriminator receives the data deceived by the constructor. As this process is repeated, counterfeit money becomes more sophisticated, making it possible to produce more and more realistic false data.[4]

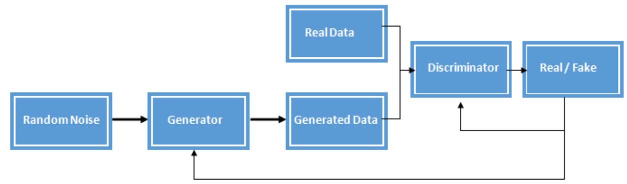

Figure 1. Generative Adversarial Network (GAN) framework

2. Related Work

GAN is mainly used for image generation. Learn the real image to produce a false image. In 2017, Nvidia showed the technology to learn the photos of 200,000 celebrities and create infinite photos of non-existent people. It is difficult for the human eye to determine whether it is a real person or a virtual person. In the past, it was faster and easier for professionals to work with Photoshop and more. Nvidia's paper shows that not only humans, but also things like bedrooms, pots, sofas, buses, etc., can make AI come true. In addition, using GAN, it is possible to create a product design clan with a simple sketch or make it look like a famous artist's painting. You can also restore damaged images, such as making low-resolution photos high resolution [4] through the architecture shown in Fig. 2.

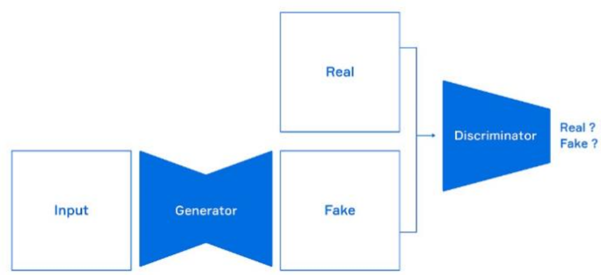

Figure2.Learning Process Principles [4] [5]

Many works has been done in recovering old paintings. Antique works are in great demand for commercial purpose as well as research purpose. There is a much need to restore old documents as well.

Raymond A. Yeh et al. proposed Semantic Image Inpainting with Deep Generative Models [6]. In their work VariationalAutoencoders (VAEs) is an emerging approach in today's deep learning methods for learning intricate distributions in an unsupervised learning method. In the paper Auto-Encoding Variational Bayes, the author proposed avariational inference and learning algorithm that can handle large datasets. They have provided many algorithms to prove the importance but it does not cover the variations on large datasets.

Xintao Wang et al. in their paper ESRGAN: enhanced super-resolution generative adversarial networks, they developed enhanced SRGAN to improve the network architecture, adversarial loss and perceptual loss of general purpose SRGAN [8]. They removed all BN layers from the generator and replaced the original block to residual-in-residual dense block that has multilevel residual network. Author improved the discriminator based on relativistic GAN.

Jiahui Yu et al. proposed a free form image inpainting using gated convolution [9]. In their work convolution layers with filters are applied to input feature maps to generate an output feature map. Free form mask generation has also been referred to in this paper to improve image inpainting process.

Guilin Liu et al. proposed image inpaiting for irregular holessing partial convolutions [10]. They used stacked partial convolution operations and mask updating steps to perform image inpainting. Their result across all different hole-to-image area ratios are shown in Table 3.

Table 3. Comparison results with various methods [10]

	[0.01,0.1]		(0.1,0.2]		(0.2,0.3]		(0.3,0.4]		(0.4,0.5]		(0.5,0.6]	
	N	B	N	B	N	B	N	B	N	B	N	B
ℓ_1(PM)(%)	**0.45**	**0.42**	1.25	1.16	2.28	2.07	3.52	3.17	4.77	4.27	6.98	6.34
ℓ_1(GL)(%)	1.39	1.53	3.01	3.22	4.51	5.00	6.05	6.77	7.34	8.20	8.60	9.78
ℓ_1(GnIpt)(%)	0.78	0.88	1.98	2.09	3.34	3.72	4.98	5.50	6.51	7.13	8.33	9.19
ℓ_1(Conv)(%)	0.52	0.50	1.26	1.17	2.20	2.01	3.37	3.03	4.58	4.10	6.66	6.01
ℓ_1(PConv)(%)	0.49	0.47	**1.18**	**1.09**	**2.07**	**1.88**	**3.19**	**2.84**	**4.37**	**3.85**	**6.45**	**5.72**
PSNR(PM)	32.97	33.68	26.87	27.51	23.70	24.35	21.27	22.05	19.70	20.58	17.60	18.22
PSNR(GL)	30.17	29.74	23.87	23.83	20.92	20.73	18.80	18.61	17.60	17.38	16.90	16.37
PSNR(GnIpt)	29.07	28.38	23.20	22.86	20.58	19.86	18.53	17.85	17.31	16.68	16.24	15.52
PSNR(Conv)	33.21	33.79	27.30	27.89	24.23	24.90	21.79	22.60	20.20	21.13	**18.24**	18.94
PSNR(PConv)	**33.75**	**34.34**	**27.71**	**28.32**	**24.54**	**25.25**	**22.01**	**22.89**	**20.34**	**21.38**	18.21	**19.04**
SSIM(PM)	**0.946**	**0.947**	0.861	0.865	0.763	0.768	0.666	0.675	0.568	0.579	0.459	0.472
SSIM(GL)	0.929	0.923	0.831	0.829	0.732	0.721	0.638	0.627	0.543	0.533	0.446	0.440
SSIM(GnIpt)	0.940	0.938	0.855	0.855	0.760	0.758	0.666	0.666	0.569	0.570	0.465	0.470
SSIM(Conv)	0.943	0.943	0.862	0.865	0.769	0.772	0.674	0.682	0.576	0.587	0.463	0.478
SSIM(PConv)	**0.946**	0.945	**0.867**	**0.870**	**0.775**	**0.779**	**0.681**	**0.689**	**0.583**	**0.595**	**0.468**	**0.484**
IScore(PM)	0.090	0.058	0.307	0.204	0.766	0.465	1.551	0.921	2.724	1.422	4.075	2.226
IScore(GL)	0.183	0.112	0.619	0.464	1.607	1.046	2.774	1.941	3.920	2.825	4.877	3.362
IScore(GnIpt)	0.127	0.088	0.396	0.307	0.978	0.621	1.757	1.126	2.759	1.801	3.967	2.525
IScore(Conv)	0.068	0.041	0.228	0.149	0.603	0.366	1.264	0.731	2.368	1.189	4.162	2.224
IScore(PConv)	**0.051**	**0.032**	**0.163**	**0.109**	**0.446**	**0.270**	**0.954**	**0.565**	**1.881**	**0.838**	**3.603**	**1.588**

3. Methodology and Results

In order to utilize a hostile generating neural network, a discriminating network is required along with a generating network. In a typical hostile generating neural network, a discriminating network is a single network that determines whether the generated image is true or false. However, the simple determination of the generated image does not have a contextual consideration, which makes it difficult to generate a natural image. The learning of the proposed network consists of the learning of the generating network and the discriminating network. In the learning process, these networks can be learned through organic action, so to learn the network, learn by inputting 1) original image, 2) image damaged by Mosaic transformation, and 3) damaged area mask.[1]

As documents get old, their quality also degrades. In our work, we have used Super Resolution Generative Adversarial Network (SRGAN) as shown in Fig. 3 to generate high quality images as the initial step. So, the first step is to enhance the quality of the old and low quality image. After that, we implemented semantic image inpainting using a generative model. We have collected old document images from google and also manually from libraries. Our dataset contain 10,000 images with which SRGAN has been trained.

In super resolution GAN a low resolution image is passed through the generator, which then produces a super resolution image. The super resolution and high resolution image is then passed through the discriminator which distinguishes both the image and back-propagates the adversarial network to train the generator and the discriminator [6]. SRGAN consists of convolutional layers, parameterized rectified linear unit and batch normalization.

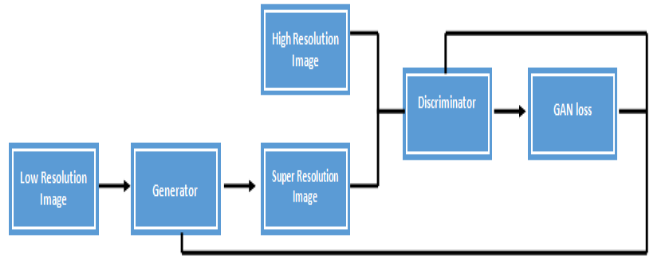

Figure3. Super Resolution Generative Adversarial Network (SRGAN) framework

For the generator of SRGAN, the loss function consists of the content and adversarial loss, while to train the discriminator, the general GAN discriminator is used for loss function. Fig. 4, Fig. 5, Fig. 6, Fig. 7, Fig. 8 and Fig. 9 are some sample results of using SRGAN for obtaining super quality images or we can say for enhancing the image quality.

Semantic image inpainting is used to fill any missing region in an image by using a generative model that trains both the generator and the discriminator with uncorrupted images so that it can recover the corrupted image closest to the original image with high quality [7]. Our proposed methodology result has been obtained by marking the damages, extracting the threshold binary of the image, generating the mask and restoring the image.

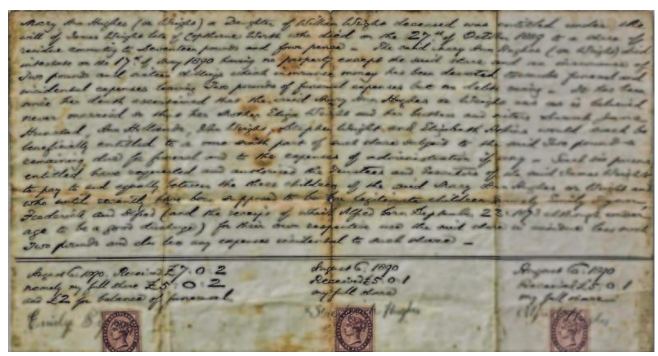

Figure4. Original image

Figure5.Result image by SRGAN

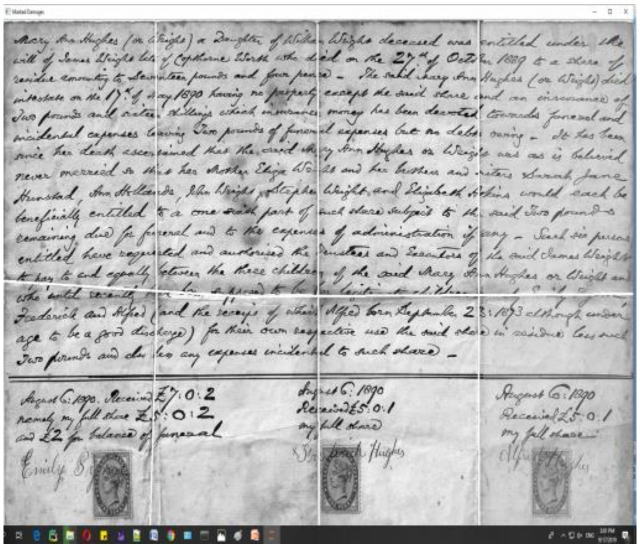

Figure6.Damages Marked

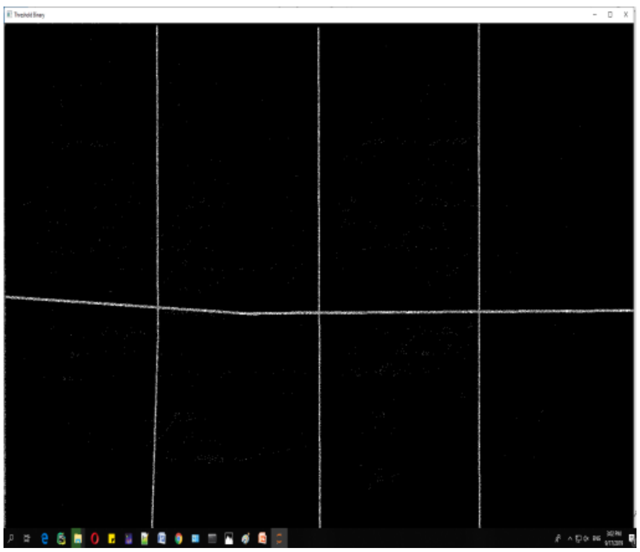

Figure7.Image of Mask

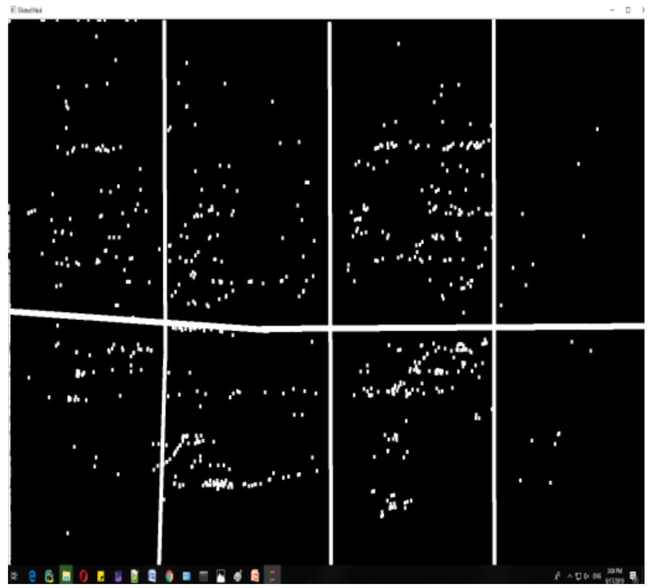

Figure8.Thresholded Binary

Figure9.Restored Image

4. CONCLUSION

In this paper, we propose a method to reconstruct an image using contextual information of surrounding images, using a deep neural network and a hostile generation network. Also, in reconstructing these images, we proposed a method that can reconstruct a more realistic image by reconstructing the detailed image and texture at the same time as the natural reconstruction by adjusting the variable value of the algorithm.

We have tested super resolution generative adversarial network and semantic image inpainting process with 3,500 old documents and generated very good results, almost accurate in every cases.

In future, we plan to extend this work to improve the restoration process and increase the accuracy of the recovery of the damaged documents and also to extract every text from the restored image separately to complete the recovery process completely.

5. ACKNOWLEDGMENTS

This work was supported by 'Jeju Industry-University Convergence Foundation' funded by the Ministry of Trade, Industry, and Energy(MOTIE, Korea). [Project Name: 'Jeju Industry-University convergence Foundation / Project Number: N0002327]. Also, this research was financially supported by the Ministry of SMEs and Startups(MSS), Korea, under the “Regional Specialized Industry Development Program(R&D or non-R&D, Project Number: P0006418)” supervised by the Korea Institute for Advancement of Technology (KIAT).

6. REFERENCES

[1] ChanKyoo Moon, YoungJung Uh, Hyeran Byun, 2018.7, 503-510, The Korean Society Of Broad Engineers

[2] Cao, Y.J., Jia, L.L., Chen, Y.X., Lin, N., Yang, C., Zhang, B., Liu, Z., Li, X.X. and Dai, H.H., 2018. Recent Advances

of Generative Adversarial Networks in Computer Vision. IEEE Access, 7, pp.14985-15006.

[3] https://developers.google.com/machine-learning/gan/loss

[4] [NaverWikipedia] Genetic Hostile Neural Networks (GANs) - AI to create a fake like real (IT in terms of) https://terms.naver.com/entry.nhn?docId=5645660&cid=59088&categoryId=59096

[5] Naver Labs, GAN Learning Process Principles

[6] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z. and Shi, W., 2017. Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4681-4690).

[7] Yeh, R.A., Chen, C., Yian Lim, T., Schwing, A.G., Hasegawa-Johnson, M. and Do, M.N., 2017. Semantic image inpainting with deep generative models. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 5485-5493).

[8] Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y. and Change Loy, C., 2018. Esrgan: Enhanced super-resolution generative adversarial networks. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 0-0).

[9] Yu, J., Lin, Z., Yang, J., Shen, X., Lu, X. and Huang, T.S., 2018. Free-form image inpainting with gated convolution. arXiv preprint arXiv:1806.03589.

[10] Liu, G., Reda, F.A., Shih, K.J., Wang, T.C., Tao, A. and Catanzaro, B., 2018. Image inpainting for irregular holes using partial convolutions. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 85-100).